Session 8: Linked Repositories – Theme of the Day

Moderator: Kristen LaBonte

The Role Played by NaFIRRI Library and Data Centre in Managing Scientific Fisheries Data From Ugandan Waters

Alice Endra

Abstract:
National Fisheries Resources Research Institute (NaFIRRI) is one of the Public Agricultural Research Institutes under Uganda's National Agricultural Research Organisation (NARO). National Fisheries Resources Research Institute Information and Data Centre has been in existence for over 50 years. It is charged with not only managing library holdings but is also with managing the Institute's scientific data generated from research. The library serves as a data center, meaning that the staff also work as data librarians. It acts as a storehouse for the Institute's data, most of which is maintained in their raw form. The data date back in the 1920's when the first surveys were carried out on Lakes Victoria and Albert in Uganda. The datasets have been organized in files. The paper explains the different ways used by NaFIRRI Library to organize and manage scientific data. It will examine challenges with managing raw scientific data that the library has faced.

Keywords: Libraries, fisheries data, data management, data curation, data center, Uganda.

Introduction

National Fisheries Resources Research Institute (NaFIRRI) is one of the Public Agricultural Research Institutes under National Agricultural Research Organisation (NARO). Its mandate is to conduct basic and applied research of national and strategic importance in aquaculture, capture fisheries, water environment, socio-economic and marketing, Information communication management and emerging issues in the fisheries sector.

National Fisheries Resources Research Institute Library and Data Centre has been in existence for over 50 years. NaFIRRI Library and Data Centre is charged with not only managing library holdings but also with managing the Institute's scientific data generated from research. The library serves as a data center, meaning that the staff members also work as data librarians. It acts as a storehouse for the Institute's data, most of which are maintained in their raw form.

The data date back to the 1920's when the first surveys were carried out on Lakes Victoria and Albert in Uganda. The data sets have been organized in files.

NaFIRRI Library and Data Centre was established in 1948 to provide information to scientists in the East African region. The library after renovation in 2005 became a Library and Information Centre, meaning that it had now acquired the additional responsibility of being a data center. To date the data centre possesses historical fisheries data from the 1950's in raw form. Therefore the library:

1). Acts as an archive for the institute's data.
2). Organizes data and information for easy access and protects them from damage and loss.
3). Data and information are available on request to scientists and the various stakeholders.

The respondents interviewed admitted that they access data from the information center. The reason why the data are kept in their raw form is because researchers often request the data files and re-analyze the data to make comparisons between the past and present. This introduces the subject of data curation in libraries. According to Jahnke et al. (2012), data curation is a term defined as a set of activities that includes the preserving, maintaining, archiving, and depositing of data to keep it secure, intact, and accessible for reuse.

Heidorn (2011) states that there are a number of arguments that lead to the conclusion that libraries must curate digital data to protect and disseminate the intellectual capital of society. Curation of the data is within the libraries' mission, and libraries are among the only institutions with the capacity to curate many data types. The data are critical to the scientific and economic development of society. The reason for conducting all of this work is of course for access, use, and reuse of the data.

David Stewart (2015) argues that increased interest in scientific data and a need for data-centric services provide a host of opportunities for the library to re-establish itself as having a central role within research institutions, but the overlapping roles of competing organizations mean that other organizations can quickly stake claims in areas that the library profession may have considered theirs by right.

The library is well situated to be a key player in data management, curation, and preservation, given its extensive experience with selection, metadata, collections, institutional repositories, preservation, curation and access (Erway 2013).

The main objectives of the study were:

1. To evaluate the role of the library and data center in managing scientific fisheries data.
2. Identify the gaps in management of fisheries data.
3. Identify ways of improving the management, access and sharing of scientific fisheries data and information within the library.

Some of the Historical Data Available

1. Lake Victoria Bottom and mid water trawl data, catch assessment survey data and frame survey data.
2. Biometric data for *Lates niloticus* (Nile perch) studies.
3. Lake Wamala data.
4. Lake George data.

5. Lake Mburo data.
6. Lake Nabugabo data.
7. Lake Albert data.
8. Lake Kyoga data.
9. Gillnet data.
10. Catch composition data.
11. Water quality data.
12. Fisheries socioeconomic data.

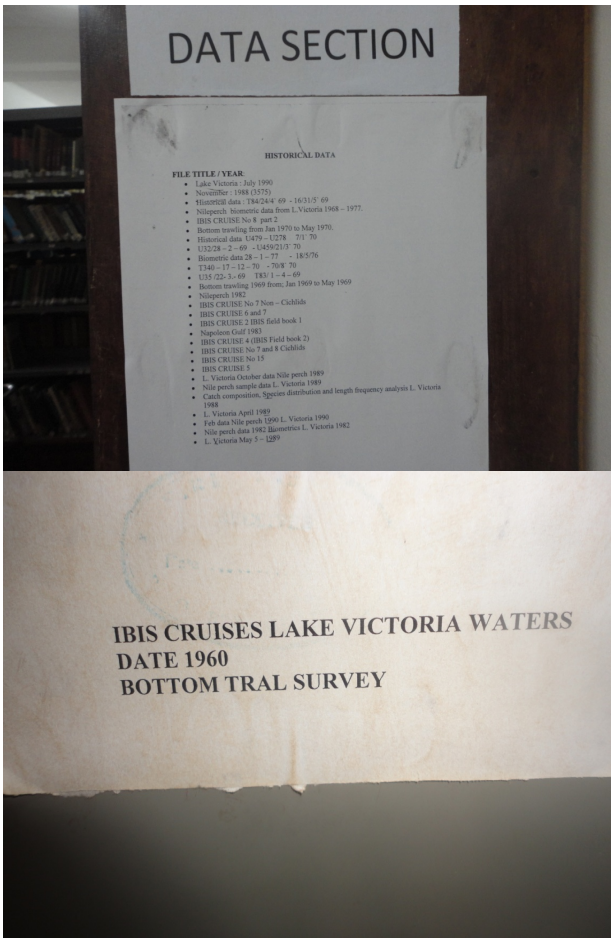

Figure 1: Part of the Historical Fisheries Data.

Role of the NaFIRRI Library in Managing Scientific Data

According to Heidorn (2011), the role of libraries is to collect, preserve, and disseminate the intellectual output of the society. This output includes books and serials as well as the digital versions of the same. The library had managed to organize historical data from 1960's to 1990's, which can be referred to by scientists.

Data have been organized:

1. According to Lakes System**s:**
    a) Lake Albert.
    b) Lake George.
    c) Lake Wamala.
    d) Lake Victoria.
    e) Lake Kyoga.
    f) Lake Bisina.

Figure 2 : Data Organized According to Lake Systems.

2. According to Bays within the different lake systems
    a) Lingira Bay.
    b) Napoleon Gulf.
    c) Buvuma Channel.
    d) Itome Bay.
    e) Ingira Bay.

3. All on Lake Victoria: Organized in files, cabins and shelves.

Figure 3: Data Displayed in Files.

Methodology Used

The author used questionnaires and interviews to get information from the various respondents.

Key Findings From the Interviews

A. How frequently do scientists visit the library?

30 respondents were interviewed on the frequency of visits to the library and data center, as seen the table below:

	Number of times	Percentage
a)	Once a week	22.7%
b)	Twice a week	27.2%
c)	Daily	22.7%
d)	Once a month	22.7%
e)	None	4.5%

Figure 4. *Number of Times Scientists Interviewed Admitted to Visiting the Library and Data Centre.*

B. Role Played by the Library and Data Centre

The library plays a role in organizing historical data for easy access and reference by scientists. 95.4% of respondents admitted that the information and data center provides the required information and data.

When asked about the types of data they access from the information and data center, scientists said that they accessed the following:

1. Fish stock data.
2. Frame survey data
3. Limnological data (aquatic environment, water quality).
4. Data and Information on aquaculture.
5. Fisheries socio-economics data and information.

C. Contribution of Library Data to Research

When asked whether the scientific fisheries data kept in the information center had contributed positively to their research Input /output and work, the majority (94%) of the respondents said that fisheries data kept in the information center contribute positively to their research work in the following ways.

1. By helping them to shape their current and future research.
2. By providing a reference for describing trends in fisheries production.
3. By providing a source of reference materials for writing scientific reports and other publications.
4. By helping them to accomplish tasks in time and also to address stakeholders' concerns and inquiries appropriately.

All the respondents said that data being kept in the information and data center had contributed positively to their research since they have been able to access historical data and information that they have used to compare with current data.

When asked whether they are happy with the way they keep data in their possession, the following responses were given**:**

1. Not happy (75%): They need training on how best to keep data so that they are easily accessible when required. They felt that data from one section might not easily be accessed by others in different sections, yet fisheries data are inter-dependent. Furthermore, data are kept in files, shelves and computer. Data stored on computers is prone to virus attacks, hence vulnerable to loss if there is no backup.
2. Are happy (25%): Because they can retrieve data anytime from anywhere.
3. Methods used in managing scientific data: When asked about the level of satisfaction with the current method of storing and managing scientific data in data center, 87% of the respondents were satisfied because historical data are centrally managed and can easily be accessed by users.

Challenges the Library Has Faced

1. Historical data available within the library are not in electronic format but in paper, which is vulnerable to damage and loss.
2. The current data are under the custody of different scientists for fear of data piracy.
3. Convincing scientists to release data so that they are stored and kept for future reference within the information and data center is still a challenge.

Gaps

There is a need to tag the data available to the publications that have been generated from the data within the library. There is also a need to digitize the historical data available within the library and make it accessible electronically.

Key Findings and Recommendations

When respondents were asked what they thought should be done to improve data storage and management within the library and data center, they made the following suggestions:

1. Need to manage data in a network environment. Although some scientists welcomed the idea of data sharing, especially within the Institute, many are reluctant to enter into any arrangement in which they would relinquish control over access to data.
2. The Library should liaise with scientists and technicians for current information /data.
3. The Centre should link with related institutions at national, regional and international levels for purposes of sharing/accessing data and information.
4. Few researchers are aware of the data services that the library provides. This calls for more sensitization.
5. Maintaining the process of digitalization of the available data and Information.
6. Installation of modern IT packages (computers) to enable computerization of the historical data available within the data center. Data storage and management systems should be upgraded to modern systems by computerizing them.
7. Respondents noted that everybody is a good custodian of data. However, it is not helpful when data are kept at individual levels until shared. They recommend the data be kept at the library and data center for further management and external sharing.
8. Strengthen security in the library and data center. This will build the confidence of researchers to release the data under their custody.
9. Networking and electronic data management should be adopted by the library and data center.
10. Acquire more soft copies of data because they are easier to store than hard copies, which need a lot of space. Important research findings should be also stored in soft copy form and not only hard copy.

Conclusion

Despite the challenges, NaFIRRI Library and Data Centre has tried to ensure safe storage and access of data and information. We look forward to digitization of the data and sharing them through different networks. In the area of privacy and data access control, additional tools will be developed to manage confidential data and provide the necessary security. Many researchers expressed concerns surrounding the ethical reuse of research data.

References

Erway, R. 2013. *Starting the conversation: University-wide research data management policy*. Dublin, OH: OCLC Research. Available at: http://www.oclc.org/content/dam/research/publications/ library/2013/2013-08.pdf.

Gray,J., David T. et al. 2011 *Scientific data management in the coming decade*. Available at: http://research.microsoft.com/pubs/64537/tr-2005-10.pdf.

Heidorn, P. B. 2011. The emerging role of libraries in data curation and e-science. *Journal of Library Administration* 51: 662–672. DOI: 10.1080/01930826.2011.601269.

Jahnke, L., Asher, A., Spencer, A. A., Keralis, D. C., Henry, C. 2012. *The problem of data*. Council on Library and Information Resources, Washington, DC. Available at http://www.clir.org/pubs/reports/pub154.

Osswald, A., Strathmann, S. 2012**.** The role of libraries in curation and preservation of research data in Germany: Findings of a survey. Session 116 - The role of libraries in data curation, access and preservation: an international perspective. Available at: http://conference.ifla.org/ifla78.

Stuart, D. (2015). Libraries could play key role in managing research data **in** Research Information: December 2014/January 2015. Available at: http://www.researchinformation.info/features/feature.php?feature_id=497.